DOCUMENT RESUME

ED 037 396 SP 003 628

AUTHOR Hanushek, Eric
TITLE The Production of Education, Teacher Quality and Efficiency.
PUB DATE Feb 70
NOTE 35p.; Paper presented at the Bureau of Educational Personnel Develcpment Conference, "How Do Teachers Make a Difference?" Washington, D.C. February 1970

EDRS PRICE EDRS Price MF-$0.25 HC-$1.85
DESCRIPTORS Board of Education Policy, Disadvantaged Youth, Elementary School Teachers, Socioeconomic Background, Students, Teacher Background, *Teacher Characteristics, *Teacher Influence, *Teacher Selection, *Teaching Quality

ABSTRACT

A study examined the educational process to identify teacher role in education and school efficiency in hiring teachers. The implicit educational model used by administrators is known: that a teacher's productivity is a function of experience and educational level. A conceptual model of the educational process was developed which states that education output, itself a multi-dimensional factor, is a function of the cumulative background influences of the individual's family, the cumulative background influences of his peers, his innate abilities, and the cumulative school inputs. Two separate analyses of elementary school processes, based on the model, were undertaken, one relying on data from the Northeast and Great Lakes of the "Equality of Educational Opportunity" survey (713 schools) and one using a sample drawn from a California school district in 1969 (2445) students). Educational output was measured only by achievement tests. Black, white, and Mexican-American students were considered separately for part of the study. Major conclusions were that (1) teachers do generally count in education; (2) schools operate quite inefficiently by hiring teacher attributes which lead to little or no achievement gains; and (3) there appears to be considerable latitude for public policy to improve our educational system. Teacher verbal ability, recentness of education, and specific socioeconomic class experience are the best measures of teacher quality contained in the data. (JS)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED037396

# THE PRODUCTION OF EDUCATION, TEACHER QUALITY AND EFFICIENCY

by

Eric Hanushek

## I. Introduction

It is currently in vogue to claim that the public education system is failing us. This is supported by a variety of evidence on incomes, racial disparities in achievement, and so forth. However, such statements by themselves are not very useful since, even if true, they provide the educational decisionmaker with no information from which to do his job better. It is simply easier to provide a balance sheet of the outputs of education than it is to provide prescriptions for action, and this fact accounts for why there has been more analysis of the results of education than of methods of improving education.

Hopes for improving public education in the U.S. depend upon our learning from past experiences. We must be able to assimilate the results of past educational programs and past instruction. However, the complexities of education make this assimilation very difficult. School administrators are often good at making judgments about very specific aspects of education. For example, a principal often can make a good judgment about which teachers are getting results and which are not. Yet, at the same time he has difficulty in pinpointing the characteristics

I am indebted to John Jackson for many helpful suggestions.

SP003628

which lead to "getting results." He will often conclude that it's all in the individual. But, if this is truly the case, we have little hope for improving public education. In order to improve our educational system we must be able to make some generalizations about characteristics of teachers which are more or less favorable to education.

This paper looks at the educational process with the aim of identifying the role of teachers in education. Moreover, since the implicit model of education used by administrators is known -- namely that a teacher's productivity is a function of experience and educational level, it is possible to make some statements about the efficiency of schools in their hiring of teachers. After sketching a general model of the production of education, the paper presents two separate attempts at estimating models of education. The first relies upon the data from Equality of Educational Opportunity [EEO];* the second uses a new sample collected from a California school system during the summer of 1969. From these analyses it is concluded that: 1) teachers do generally count in education; 2) schools now operate quite inefficiently; and 3) there appears to be considerable latitude for public policy to improve our educational system.

---

*James S. Coleman, et al. Equality of Educational Opportunity (Washington, D.C.: Government Printing Office, 1966), commonly known as the Coleman Report.

## II. Conceptual Model

It is not possible to look at the role of teachers in education in isolation. Instead, one must consider all of the factors that enter into educational process and how they interact with one another. Thus this study of the effects of teachers on the education of children rightfully starts with a discussion of a larger model of the educational process and the various factors that enter into it. After presenting an abstract model of the educational process, this section considers specific measurement of the various inputs to the educational process and the outputs of the educational process. If one can identify and measure the effects of schools and teachers on education of individual children, then one can make some statements on how best to organize the school to provide the most educational output.

The basic model of the educational process can be depicted by an equation such as Equation 1.

(1) $\underline{A}_{it} = f(\underline{B}_i^{(t)}, \underline{P}_i^{(t)}, \underline{I}_i, \underline{S}_i^{(t)})$ where

$\underline{A}_{it}$ = vector of educational outputs of the $i^{th}$ student at time t

$\underline{B}_i^{(t)}$ = vector of family inputs to education of $i^{th}$ student at cumulative time t

$\underline{P}_i^{(t)}$ = vector of peer influences of $i^{th}$ student cumulative to time t

$\underline{I}_i$ = vector of innate endowments of $i^{th}$ student

$\underline{S}_i^{(t)}$ = vector of school inputs to $i^{th}$ student cumulative to time t

This model simply states that educational output ($\underline{A}_{it}$), itself a multi-dimensional factor, is a function of the cumulative background influences of the individual's family ($\underline{B}_i(t)$), of the cumulative influences of his peers ($\underline{P}_i(t)$), of his innate abilities ($\underline{I}_i$) and of the cumulative school inputs ($\underline{S}_i(t)$). While this abstract model is not very operational, it does provide a framework for discussion of models of the educational process which can be tested empirically.

Specific measures of each of the inputs listed in Equation 1 are derived from a combination of past work in the field, theoretical considerations, and sheer data availability. For instance, one can think of many measures of the output of the educational process. It would be possible to use standardized test scores, juvenile delinquincy rates, future income streams, or level of education completed. However, for any given sample of data one is usually hard pressed to find more than one of these specific measures. While theoretically one thinks of schools producing several different outputs, usually lumped under the major categories of cognitive development and socialization, the availability of data has restricted most past studies to examining a single output. Indeed, this will be the situation in the analysis that is presented in this paper. This paper concentrates entirely on an analysis of cognitive development as reflected in scores on standardized ability and achievement test scores.* It is

---

*Two different tests are used in the course of the

believed that these scores represent differences which are valued by society.*

The inputs are also subject to many of the same considerations as the measure of output. There is no firm theoretical basis for choosing inputs. Likewise, there is often a lack of desired data. Each input vector will be discussed in turn.

Families contribute to the education of children in many different ways. They provide basic shelter and food for the individual child. But more than that, they provide models of verbal structure, examples of problem solving, and a basic set of attitudes to the individual child. To measure each of these concepts explicitly would be a very difficult task, but for our purposes this is not really necessary. It is widely accepted that the relevant educational inputs are highly correlated with the socio-economic status of the family. Thus one can indirectly include the effects of each of these individual family inputs in the

---

analysis: 1) Educational Testing Service's School and College Ability Test (SCAT) for verbal ability in grade 6, and 2) Stanford Achievement Test for reading in grade 3.

*There is scattered evidence on this in W. Lee Hansen, Burton A. Weisbrod and William J. Scanlon, "Determinants of Earnings of Low Achievers: Does Schooling Really Count, Even for Them?", mimeo, Institute for Research on Poverty, University of Wisconsin, February 1969; Burton A. Weisbrod and Peter Karpoff, "Monetary Returns to College Education, Student Ability and College Quality," The Review of Economics and Statistics, November 1968; and Randall D. Weiss, "The Effects of Education on the Earnings of Blacks and Whites," Discussion Paper No. 44, Program on Regional and Urban Economics, Harvard University, April 1969.

educational process by including a set of measures of socio-economic status. These measures include parents' educations, goods in the home, family size and father's occupation.

Peer groups provide many of the same inputs that the families provide. The individual child's peer groups would include his friends both inside and outside of school. To be precise, one would want to know exactly which individuals were friends or tended to interact with each other, but collecting this kind of information on a very large scale would be prohibitively expensive. In this case, it seems acceptable to aggregate all classmates of the individual in the classroom or school and take that as the peer groups. In measuring the interactions of individual children one can use the same proxies for peers that are used in the case of the individual's family, that is, use socio-economic status as a proxy for the types of interaction which exist among friends. Thus for peer groups we would want to take aggregates of the individual family background measures.

Innate abilities is probably the most difficult concept to measure in the whole model. In fact, it is not well understood how innate abilities enter into the educational process, and there exists considerable controversy over the role of innate ability in education. The only consensus which appears to exist in the area is that common IQ scores do not do an adequate job of measuring innate abilities. All is not lost, however, when innate abilities cannot be measured directly. In particular, under a set of

plausible assumptions (which will be detailed in the empirical section) it is possible to circumvent the most serious problems.

School influences are the focus of this study and will be discussed in more detail than the other inputs. The hypotheses to be analyzed actually are quite simple and straightforward. It is surprising how little is actually known about the ways in which schools and teachers affect education. This largely results from a fixation on inputs to education rather than outputs. However, one can impute a set of hypotheses about teacher effects from the behavior of schools. In particular, schools base pay schedules on teaching experience and educational levels. Thus, they must believe that increased experience and further schooling have a positive relationship to educational output. These provide two central hypotheses in the study of the educational process.

Other hypotheses can also be found in the actions of school administrators. A frequent compensatory education plan is the reduction of class size. Since this is a very expensive undertaking, the presumed benefits (increased outputs) must be great. Also there are a large number of people who argue that some forms of student distributions in the schools and classrooms (e.g., ability tracking or racial and social integration) have a beneficial effect on education.* All of these are testable hypotheses about the relationship between school inputs and achievement.

---

*Cf. U.S. Commission on Civil Rights, Racial Isolation in the Public Schools (Washington, D.C.: Government Printing Office, 1967), Chapter III.

Further, in recent literature (particularly Equality of Educational Opportunity /EEO/ there is a suggestion that one can measure other dimensions of teacher and school quality. These include attitudes of teachers and administrators, verbal facility (and perhaps general ability) of teachers, quality of physical plant, quality of teacher education, background of teachers, and more.

Together the above forms the rudiments for a testable model of the educational process. While some modifications are required because of data limitations, this basic structure will hold in the empirical section.

## III. Empirical Analysis

Two separate analyses of the educational process in elementary schools area have been undertaken in this paper. The first relies upon the data for the Northeast and Great Lakes of Equality of Educational Opportunity. The second uses a sample drawn from a California school district during 1969. Each of these analyses will be described separately, and then they will be compared for consistency and conclusions.

### Multisystem School Analysis*

The well known report Equality of Educational Opportunity assembled the best data bank on public education to date. This

---

*This section relies heavily on analysis presented in more detail in Eric Hanushek, "The Education of Negroes and Whites" (Unpublished PhD dissertation, Massachusetts Institute of Technology, 1968).

1965 survey collected a wealth of data pertaining to students, schools and the outcomes of education. A reanalysis of these data comprise the first section of applications of the basic educational model.*

The survey collected data on some 570,000 students and 67,000 teachers across the country. It was a purely cross sectional survey of students in grades 1, 3, 6, 9, and 12. Minorities were intentionally overrepresented in the sample.

The student information included a set of standardized test scores (verbal ability, nonverbal ability, reading achievement, and mathematics achievements) and questionnaire responses to both objective questions about the students' background and subjective questions about the students' attitudes toward school and society and the parents' attitudes about similar issues.

The teachers in the sampled schools completed a questionnaire concerning objective background characteristics (education, family background, experience, etc.) and subjective characteristics (attitudes toward students, minorities, compensatory education, etc.). They also completed a simple verbal facility test.

Finally, principals and school superintendents supplied information on general school characteristics, curricula, and

---

*The shortcomings of the analysis in Equality of Educational Opportunity which suggest a reanalysis would be valuable are discussed elsewhere. Cf. Eric Hanushek and John Kain, "On the Value of Equality of Educational Opportunity as a Guide to Public Policy," Discussion Paper No. 36, Program on Regional and Urban Economics, Harvard University, 1968.

their personal backgrounds and attitudes.

In using these data to test the model of the educational process, two factors are immediately evident. The data do not relate school and teacher inputs to individual students. In no place is there any information on specific inputs received or available to an individual student. One only knows what school averages look like. Therefore, there would be considerable error in the school input variables if one attempted to estimate a model for individuals like Equation 1. Secondly, there is no measure of innate abilities in the model.

The first problem, the inability to estimate models for individual students, is overcome by looking at total school models. Instead of using the achievement of individual students as the output of the educational process, students are aggregated across schools so that average scores for a given grade represent the output. At the same time, inputs are aggregated across the school so that average background characteristics and average school characteristics form the inputs. This tends to minimize the data problems introduced by incompatibility of student and school data.

One obvious loss from this aggregation is the influence of peers on students. It is no longer possible to differentiate between family backgrounds (in aggregated form) and peer influences. (One crude peer effect can be analyzed. This is the effects of one racial group on others. However, this becomes tricky to interpret because of the intertwined and competing hypotheses

involved in the racial influence variables.)

Innate abilities are not handled as neatly. There is no direct measure. However, at least for whites, it is reasonable to assume that this factor is fairly well captured in the family background variables. This is the case if innate abilities tend to be hereditary and if social mobility is highly correlated with ability.* For blacks, where the parent to son correlations of SES are not nearly as pronounced, this logic is more strained.** The principal problem arising from lack of measure of initial endowments is biased statistical results. But bias only arises when the excluded variable (innate abilities) is not independent from the included inputs. Thus, even in the black case, severe problems at least at the school level do not arise unless there is a mechanism which leads to the correlation of innate abilities and specific school resources. For the purposes of analyzing school and teacher influences this omission, then, does not seem too damaging. Note, however, that this factor further complicates the family background factors. Those who would attempt to derive policy implications from the background portions of the model are warned again of the extremely complicated nature of that set of inputs.

The specific school analysis undertaken involved estimating

---

*Peter M. Blau and Otis D. Duncan, The American Occupational Structure (New York: John Wiley and Sons, 1967).

*See The American Occupational Structure.

separate black and white models. Separate models were estimated for two reasons. First, since many of the inputs -- particularly the background factors -- are measured by social class proxies, there is no reason to assume that these nominal measures imply the same behavioral content. Secondly, there is no reason to assume that the educational process is the same across racial lines. In fact many people maintain strongly that differences do exist.

The analysis is concentrated upon the sixth grade students in the sample. This choice was the result of two factors. The inability to include historical information due to the cross sectional survey with little data on the past indicated that data from earlier schooling with less chance of moves, changes in status, etc. introducing error would be superior. However, there was a trade-off here because the students supplied all of the information on their background (no consultation with parents); going back to the first and third grades would introduce a different type of data error. The desirability of using elementary schools for the analysis is immediately obvious. The generally simpler school organization, the more standardized curricula, and the more homogeneous size make elementary schools much more attractive for modelling than intermediate or high schools.

The samples used for the analysis included all urban elementary schools from the Northeast and Great Lakes regions of the Equality of Educational Opportunity survey that had at least five white or black sixth graders. This yielded 471 schools with five or more white students and 242

schools with five or more blacks. In both samples the racial mix contains observations across the whole spectrum from less than 5 percent of the opposite race to over 95 percent, although both samples are heavily represented by highly segregated schools.

Results -- multisystem, school analysis

Models of education for whites and blacks were estimated using regression techniques.* In both cases a multiplicative (log-log) functional form proved superior to a linear form. Thus, the estimated coefficients can be interpreted as elasticities.** Three separate measures of teacher quality proved significant in the models: teacher experience, teacher verbal facility test scores and the percent of students with a nonwhite teacher during the previous year. The effects of teachers on the production of verbal achievement is presented in Table 1 along with the means and standard deviations. The complete models are found in Appendix A. Since the focus of our attention is on the effects of teachers, only teacher effects are shown in Table 1 even though the estimates were derived from a larger model. Suffice it to say here that the background variables appear to do a good job of measuring home and peer influences on education. Further, the estimated effects of

---

*Because of the heteroscedastic errors introduced by using school observations, weighted regression techniques were used to improve the efficiency of the estimators. See "The Education of Negroes and Whites," Appendix A.

**An elasticity presents the percentage change in verbal achievement that will result from a one percent change in the given input. Mathematically,

$$\text{elasticity} = \frac{\% \text{ change in verbal score}}{\% \text{ change in input value.}}$$

TABLE 1

TEACHER EFFECTS ON VERBAL ACHIEVEMENT, MEANS AND STANDARD DEVIATIONS

| Variable | Elasticity | Mean | Stnd. Dev. |
|---|---|---|---|
| WHITE MODEL | | | |
| Teacher Experience (years) | .020 | 11.9 | 4.6 |
| Teacher Test Score | .117 | 24.8 | 1.4 |
| % Students with Nonwhite Teacher Last Year | -.024 | 13.4 | 16.0 |
| BLACK MODEL | | | |
| Teacher Experience (years) | .045 | 11.3 | 4.0 |
| Teacher Test Score | .178 | 24.0 | 1.8 |
| % Students with Nonwhite Teacher Last Year | -.026 | 44.7 | 19.4 |

Complete model: Verbal = f(goods in home, father's education, family size, attitudes, central city, racial composition, and teachers)

teacher inputs seem to be invariant to the precise formulation of background factors and to the inclusion or exclusion of the attitudinal variables.

Since the school influences in the two models appear quite similar, it is possible to discuss both models at the same time. One of the more interesting features of the models is that only one factor which is explicitly purchased by schools affects achievement; this is teacher experience. Further, the small coefficients

indicate that experience does not have an overwhelming effect on achievement. The existence of "seniority rights" in school selection suggests an upward bias as school achievement could well influence selection by teachers. However, indirect evidence of the insignificance of direct attitude variables about school selection by the teachers indicates that this variable is chiefly a "pure" experience measure. It is somewhat surprising that the elasticity is constant across the whole range of experience, although tests for differences in different ranges proved insignificant.

The teacher verbal test score represents the best measure of teacher quality contained in the data. This provides a method of making standardized comparisons across teachers but is a still crude measure of teacher quality. It gives some measure of the technical competence of the teaching staff in one particular dimension -- verbal ability, and it probably acts as a partial proxy for general intelligence. Nevertheless, there are many other dimensions of teaching, e.g., rapport with the class, empathy, warmth, knowledge of subject matter, which are valuable in teaching but not included in this measure.* Given these shortcomings, the magnitude of the effect is significant. The elasticity of .12 (.18) for such a poorly measured indicator of teacher quality provides considerable encouragement in the ability

---

*The narrowness of this quality measure is further attested to by similar analysis of the production of mathematics achievement test scores. In those models the elasticity drops to .09 and the t-ratio goes to 1.3. This indicates a more narrow technical competence interpretation.

of schools to affect children. Table 2 indicates the small variation in this measure; the standard deviation for whites equals only 1.4 with a mean of 24.8 and a maximum score of 30 with a black sample mean approximately one point less. Nevertheless, there are wide fluctuations of scores even within cities. Within one sampled city, there were differences of 40 percent between the best and worst schools.* Switching the teacher staffs would result in a five to seven percent increase in average achievement.

The final teacher quality measure is the percentage of sixth graders who had a nonwhite teacher during the last year. This is interpreted as a measure of part of the teacher quality distribution, i.e., the lower end of the distribution. This interpretation arises from our knowledge of the education provided to blacks. Many studies, including a survey of colleges presented in Equality of Educational Opportunity, show the general quality gap between Negroes and whites who go into teaching.** This is not particularly surprising given that blacks are given inferior elementary and secondary school education and then proceed to segregated colleges which tend to widen the educational gap (by race).***

---

*The other teacher variables in these schools were roughly equal.

**EEO, Chapter IV and James A. Davis, Undergraduate Career Decisions (Chicago: Aldine Publishing Co., 1965).

***EEO, Table 3.121.1.

Before discussing the larger implications of these results, it is useful to digress for a moment and discuss some of the school factors which proved insignificant in modelling the educational process. These include teacher degree level, sex, age, teaching certificates, attitudes toward teaching and the students, measures of teacher background, and class size. Certainly, there are considerable measurement errors in each and these errors will affect the significance of the various factors. However, none seems to exert a strong influence on achievement.

A few general conclusions arise from this analysis. First, the general low effect of purchased aspects of teachers (advanced education and experience) indicates that schools are acting inefficiently. Since school systems pay handsome bonuses for these attributes, it is only economical to have people with advanced degrees if they contribute a proportionately higher amount to achievement. This does not appear to be the case.

However, these models do not support the contention that schools do not count. To the contrary, they imply that higher quality teachers do produce higher levels of achievement. Further, given the general problem of measurement errors in the data and the crudeness of the variables, the coefficients tend to be underestimated or biased downward.* Looking at Table 1, there is also the distinct impression that teacher quality impacts more on

---

*See J. Johnston, Econometric Methods (New York: McGraw-Hill Book Co., 1963), pp. 148-150.

blacks than on whites. While differences in the coefficients are small, they are consistent. If in fact this is the case, it indicates that schools can increase educational achievement for whites and blacks by allowing for these differences in the educational process. For example, it would be able to increase black achievement without changing white achievement by shuffling teachers with more experience into predominantly black classrooms (and possibly compensating predominantly white classrooms with more verbal teachers).

It is unreasonable to push these models too hard. They make two essential points. First, teachers do appear to matter. Better teachers (better here in a very limited way) achieve better results. Second, schools appear to be inefficient. They appear to be hiring the wrong things.*

Single System, Individual Student Analysis**

A similar type of analysis was carried out with a different set of data which allowed a more accurate measure of the teacher inputs received by each child. In particular, individual students were matched with individual teachers. This allowed for an

---

*This should be qualified somewhat. Even with fixed salary schedules, Henry Levin in Recruiting Teachers for Large City Schools (forthcoming) shows that it is possible to estimate supply functions for other characteristics -- particularly things like teacher verbal test scores.

**The analysis presented in this section is part of an ongoing study of education sponsored by The RAND Corporation. However, this should not be taken to represent the official views of The RAND Corp.

historical element to be introduced by matching with past teachers and alleviated the need to estimate school production functions. Thus, the data came much closer to the conceptual model of Equation 1.

The basic sample of data was drawn from a large school system in California during the summer of 1969. All children in the third grade during the school year 1968-1969 were initially included in the sample. For these 2445 students, information on family background, scores on the Stanford Achievement Tests, and names of teachers was abstracted from cumulative records. At the same time, all kindergarten through third grade teachers currently in the system were surveyed for information fairly similar to that contained in Equality of Educational Opportunity. Information was collected on teacher backgrounds, attitudes and specific aspects of schooling. An attempt was made to ascertain their use of time, i.e., the division in the classroom between instructional efforts, disciplinary efforts and administration. Also, a verbal facility test was given each teacher.* The sample used for this analysis was developed by applying two criteria to this group of all third graders. First, individuals were eliminated from the sample if data were not available on both their second and third grade teachers. Second, students were eliminated if both first

---

*Edgar F. Borgatta and Raymond J. Corsini, Quick Word Test: Level 2 (New York: Harcourt, Brace and World, Inc., 1964). This test appears to be superior to the test in Equality of Educational Opportunity as it appears to give better discrimination among teachers. One complaint voiced about the EEO test is that it was too easy.

and third grade achievement test scores were not available. When these criteria were applied, a total of 1061 students was left in the sample.

This sample allows another method of dealing with the problem of initial endowments. In particular, since there is a measure of previous test scores, it is possible to restrict the analysis entirely to one period of schooling by including the previous score for an individual as an input into the process. In this matter all of the level determining aspects of innate abilities can be eliminated. This seems to go a long way toward minimizing any biases arising from this missing information.

Looking at one school district has both advantages and disadvantages. Many hard-to-measure attributes of a school such as curriculum, school organization, community attitudes, etc. are automatically taken care of by looking at one school system. Thus, potential biases from community or system specific variables which cannot or are not measured are eliminated in such a sample. However, the same arguments can be turned around in the other direction. By looking at only one system it is difficult to make generalizations about behavior in other systems located in different regions and having different types of organization. If system specific attributes are very important, it might not be possible to apply estimated models to other systems. This implies that the previous section's analysis and the analysis in this section are very much complements of each other. Each has

weaknesses, but consistency in the different samples would strengthen the results considerably.

Empirical Results

For analytic purposes the sample was divided into sub-samples. First, whites and Mexican-Americans (the only minority group represented in the system) were separated. This follows the reasoning given for looking at whites and blacks separately. The nominal values of the proxies for background inputs do not necessarily have the same meaning for the two groups, and there is no reason to insist on the same model of the educational process for both groups. Further, the ethnic samples were divided on occupational grounds -- fathers in manual or blue collar occupations and non-manual or white collar occupations. This left three samples: white, manual occupation (n = 515); white, non-manual occupation (n = 323); and Mexican-American, manual occupation (n = 140).*

The first step in analyzing the data was to estimate third grade achievement ($A_3$) models using only the teacher inputs which are purchased by the system to represent school effects. Two linear regression models were estimated (one using

---

*These samples are not exhaustive. Children with only mothers or no occupation reported for fathers were not included. For whites, these groups totaled 36 students; for Mexican-Americans, these groups plus the non-manual occupation group totaled 47. These samples were too small to study separately, and, thus, they were ignored.

first grade achievement as an input, the other not using it). The "pay parameters" of years of teaching experience, possession of a master's degree (=1) or not (=0), and the number of college units beyond the highest degree represented the school inputs in the models. These attributes pertained to the specific second and third grade teachers for each student.

As Table 2 and Table 3 ably demonstrate, there is a general lack of statistical significance of these factors.* Only four of eighteen coefficients in the gross output case have significant t values; none in the value added case have significant t values. Further, of the significant coefficients, one has the wrong (unexpected) sign. The other three coefficients apply to the number of units beyond the highest degree and, thus, have no meaning when degree level (MASTER) is not included in the model (or has an insignificant coefficient). The implication is immediately obvious -- the things that schools are buying do not appear to be valuable in the educational process.

However, the above results give minimum guidance to an administrator. While they indicate what he should not do, they give a very imperfect picture of what he should do. For his purposes we wish to identify what attributes of teachers do seem to count. That is the emphasis of the remainder of this section.

Separate models using different measures of teacher characteristics were again estimated for white, white collar;

---

*When $|t| < 1.96$, it is not possible to reject the hypothesis that the coefficient equals zero at the 5% level.

TABLE 2: SIGNIFICANCE OF TEACHER EFFECTS (Gross output)

$A_3$ = f(sex, income, siblings, no. absences, % Mexican-American, aver. income in school, $EXPER_3$, $MASTER_3$, $UNITS_3$, $EXPER_2$, $MASTER_2$, $UNITS_2$)

| | t statistics | | |
|---|---|---|---|
| | White Manual | White Nonmanual | Mex-Amer Manual |
| $EXPER_3$ | .74 | 2.74 | -.04 |
| $MASTER_3$ | .89 | -2.69 | -.47 |
| $UNITS_3$ | 2.04 | .21 | 1.09 |
| $EXPER_2$ | -1.39 | -.55 | .77 |
| $MASTER_2$ | 1.45 | -.15 | -.42 |
| $UNITS_2$ | 2.26 | 2.93 | -.34 |

TABLE 3: SIGNIFICANCE OF TEACHER EFFECTS (Value Added)

$A_3 = f( * ) + A_1$

| | t statistics | | |
|---|---|---|---|
| | White Manual | White Nonmanual | Mex-Amer Manual |
| $EXPER_3$ | .56 | 1.69 | -.45 |
| $MASTER_3$ | .18 | -1.91 | .59 |
| $UNITS_3$ | .94 | 1.05 | 1.77 |
| $EXPER_2$ | -.61 | .30 | 1.31 |
| $MASTER_2$ | 1.94 | .60 | -.00 |
| $UNITS_2$ | .31 | -.06 | -1.60 |

for white, blue collar; and for Mexican-American, blue collar. The results for these groups were quite different. Teacher effects do not appear to be consistent across the three groups.

## White Manual

The white, manual occupation model comes closest to the previous school models. Equation 2 displays the model of the production of Stanford Achievement Test (Reading) scores estimated for 515 third graders. Variable definitions, means and standard deviations are found in Table 4. Third grade achievement is a function of the starting point (first grade achievement, $A_1$), sex (F), grade repeats (R), and a set of teacher inputs.

$$(2)\quad A_3 = \underset{}{20.8} + \underset{(2.3)}{2.81}F - \underset{(-2.8)}{6.38}R + \underset{(18.8)}{.79}A_1 - \underset{(-2.1)}{.07}D + \underset{(2.4)}{.09}T_3 - \underset{(-1.5)}{.57}Y_3$$

$$+ \underset{(1.9)}{.06}T_2 - \underset{(-2.9)}{.68}Y_2 \qquad R^2 = .51 \qquad SE = 13.5$$

Again, the interest here centers on the teacher inputs. The variable D represents the teacher's estimate of the percentage of classroom time spent on discipline. This gives some idea of the intensity of instruction received by the individual student. As expected, this has a negative impact on achievement; as more time is spent on discipline, less is spent on instruction. This suggests that there are noticeable externalities in the classroom and that efforts to reduce discipline time in the classroom would

TABLE 4

VARIABLE DEFINITIONS, MEANS AND STANDARD DEVIATIONS -

WHITE MANUAL OCCUPATION MODEL

| Variable | Mean | Stnd. Dev. | Definition |
|---|---|---|---|
| $A_3$ | 55.74 | 19.1 | Stanford Achievement Test raw score - 3rd grade |
| F | .50 | .5 | Sex: = 1 for female = 0 for male |
| R | .08 | .3 | Repeat grade: = 1 if a grade was repeated; = 0 otherwise |
| $A_1$ | 35.17 | 15.1 | Stanford Achievement Test raw score - 1st grade |
| D | 17.93 | 18.8 | % of time spent on discipline by 3rd grade teacher |
| $T_3$ | 66.90 | 15.8 | Quick Word Test score - 3rd grade teacher |
| $Y_3$ | 1.91 | 1.6 | Years since most recent educational experience - 3rd grade teacher |
| $T_2$ | 68.41 | 19.0 | Quick Word Test score - 2nd grade teacher |
| $Y_2$ | 2.64 | 2.6 | Years since most recent educational experience - 2nd grade teacher |

have positive results on achievement. For example, the principal might assume a very high proportion of discipline chores.

Two characteristics of both the second and third grade teachers were significant. Verbal facility test scores and length

of time since most recent educational experience of the teacher proved to be important attributes affecting achievement. The third grade teacher elasticity at the point of means of .11 for T and the second grade elasticity of .07 fall in line with those from the previous school analysis. It is a little surprising, however, that the elasticities are slightly less here than in the other models. The other teacher variable, Y, indicates that recent educational experiences -- either undergraduate or graduate level -- are important. Thus, efforts to have teachers return to school during summers seem justified in terms of effects on education. The cumulative effect (master's degree and total units) is not as important as recent involvement.

There are some important policy implications surrounding the verbal test measure of teacher quality. By interchanging teachers at the top and bottom of the verbal ability scale for this system, achievement changes by .2 to .4 grade levels.* This seems quite significant at this grade level, particularly if the increasing grade level disparities hypothesized in Equality of Educational Opportunity hold true for the individuals in this sample.** Thus, teacher distribution can have a significant effect on individual children. Further, since this test has national norms, it is possible to get some idea of how the

---

*This is calculated by changing only the third grade teacher verbal score for the lower limit and both second and third for the upper limit. The scores are changed from 40 to 96 to represent the range found in the data. (Maximum score is 100.) The resulting achievement score is then converted to grade level equivalents.

**EEO, Chapter 3.

teachers being hired in this system rate when compared with other college graduates. The mean score of 68 places the teachers in this sample slightly under the median for female college graduates. Thus, this system is not being successful in attracting the best people.

White Nonmanual

The model estimated for the 323 children with white collar backgrounds (Equation 3) did not show the importance of teachers to be as high as in the blue collar white sample. Definitions, means and standard deviations are found in Table 5. Equation 3 indicates that, given the first grade achievement of the student, children with fathers in clerical occupations (C) score lower. Further, the recentness of educational experience (Y) is again a factor along with the amount of experience the teacher has had with this socio-economic level(S).

$$(3)\quad A_3 = \underset{}{35.9} + \underset{(-3.0)}{.72A_1} - \underset{(-1.9)}{5.1C} - \underset{(1.2)}{.79Y_3} + \underset{(-1.7)}{.10S_3} - \underset{(1.8)}{.66Y_2} + .20S_2$$

$$R^2 = .52 \qquad SE = 11.8$$

Each of these teacher variables is statistically less significant than the teacher variables in Equation 2. Further, the magnitudes of the coefficients suggest that teachers have less effect on these children. The elasticity at point of means

TABLE 4

VARIABLE DEFINITIONS, MEANS AND STANDARD DEVIATIONS -

WHITE NONMANUAL OCCUPATION MODEL

| Variable | Mean | Stnd. Dev. | Definition |
|---|---|---|---|
| $A_3$ | 64.82 | 16.8 | Stanford Achievement Test raw score - 3rd grade |
| $A_1$ | 42.43 | 15.8 | Stanford Achievement Test raw score - 1st grade |
| C | .19 | .4 | Clerical occupation: = 1 if father in clerical job; = 0 otherwise |
| $Y_3$ | 2.02 | 1.7 | Years since most recent educational experience - 3rd grade teacher |
| $S_3$ | 7.85 | 8.1 | Years of experience with this socio-economic level - 3rd grade teacher |
| $Y_2$ | 1.88 | 1.7 | Years since most recent educational experience - 2nd grade teacher |
| $S_2$ | 7.94 | 6.1 | Years of experience with this socio-economic level - 2nd grade teacher |

for each of the four teacher variables is less than .025. Thus, changing the input values by any reasonable amount yields a considerably smaller achievement change than was found changing teacher inputs in the sample of children in blue collar families.

## Mexican-American Manual

In looking at the 140 Mexican-American children, it was impossible to find any discernible impact of schools. The best model of the educational process for these children, Equation 4, shows that in addition to entering achievement score ($A_1$), only sex (F), grade repeated (R), and differences in family background (SS and SK) affect third grade achievement. Variable definitions, means and standard deviations are found in Table 6.

$$(4)\quad A_3 = \underset{}{14.6} + \underset{(9.7)}{97A_1} + \underset{(1.2)}{2.84F} - \underset{(-2.0)}{8.92R} + \underset{(2.7)}{8.22SK} + \underset{(2.0)}{5.96SS}$$

$$R^2 = .51 \qquad SE = 13.8$$

None of the measurable factors used in this analysis concerning teachers impacted on these children, at least in the production of reading achievement. This is a shocking result, and not without its policy implications. The system has not been able to provide the type of instruction necessary for these children. Standard teaching methods do not seem to be appropriate in this case.

## Individual Student Models

In developing each of the models a set of variables corresponding to some common hypotheses about the educational process was also examined. Consistently, the influence of peers (measured by aggregate characteristics of all third graders in

TABLE 6

VARIABLE DEFINITIONS, MEANS AND STANDARD DEVIATIONS -

MEXICAN MANUAL OCCUPATION MODEL

| Variable | Mean | Stnd. Dev. | Definition |
|---|---|---|---|
| $A_3$ | 47.61 | 19.4 | Stanford Achievement Test raw score - 3rd grade |
| $A_1$ | 28.06 | 12.5 | Stanford Achievement Test raw score - 1st grade |
| F | .54 | .5 | Sex: = 1 for female = 0 for male |
| R | .08 | .3 | Repeat grade: = 1 if a grade was repeated; = 0 otherwise |
| SK | .34 | .5 | Skilled labor: = 1 if skilled occupation; = 0 otherwise |
| SS | .38 | .5 | Semi-skilled labor: = 1 if semi-skilled; = 0 otherwise |

the 25 schools for the sample) was found to be insignificant. Peer influences were measured in a number of specific ways. Occupational distribution was depicted by percentage in nonmanual occupation and average income level; ethnic distribution by percent Mexican-American. Further, ability distribution was considered in terms of average achievement scores in the first grade. For teachers, attitudes about compensatory education and minority students proved insignificant. Teacher age, sex and undergraduate major also showed no effect. Thus, the models displayed imply

a set of other hypotheses which proved insignificant.

In terms of teachers the three models can be rank ordered. Teachers have most effect on white children from blue collar families and least effect on children from Mexican-American families. This is disappointing since Mexican-American children are worst off at the beginning of the process (first grade for this analysis). The idea of schools' equalizing initial deficits of these children is obviously not realized.

For the white population teachers obviously do count. Better teachers imply better results. However, better teachers are not measured in the direction that schools measure them by their pay schedules. Instead they are measured in terms of verbal ability, recentness of education and specific socio-economic class experience. This implies that schools are being inefficient -- for a smaller expenditure on teachers schools could reach the same level of achievement. Moreover, there are gains to be made in the school systems from changing their hiring and pay systems.

## IV. Conclusions and Implications

The two separate analyses are complements. Each individual analysis has a set of problems associated with it that tends to dilute the findings. However, taken together each appears to make up for the larger problems of the other. Thus, the sum of

the two provides a much more reliable picture of education.

Throughout the analysis there is never much question about the ability to model the general educational process, at least as seen in the elementary school. As an overall view of education the models seem to do quite well. The effects depicted are consistent with a priori views; the individual elements are statistically significant; and the general explanatory power of the models seems reasonable.

The strongest conclusion from the models is that school systems now operate quite inefficiently. They are buying the wrong attributes of teachers, i.e., attributes which lead to little or no achievement gains. However, it is more difficult to develop the positive side. There are attributes which appear to be quality related which affect achievement. Yet, they can also be interpreted as proxies for other factors. To the extent that verbal facility is just a proxy for general ability or intelligence, then it is not verbal facility which we want to purchase; it is intelligence. Once a hiring policy for verbal ability was instituted, any relationship between verbal ability and intelligence would tend to disappear or possibly reverse. Thus, these models do not provide a practical guide to the school administrator. They only say that there is something there that is desirable for teachers to have.

It is strange to find strong teacher effects for blacks and

not Mexican-Americans. This suggests that it is not just deprivation or a lower educational input from outside the school. The most plausible explanation is found in the language problem. There is no measure of the intensity of Spanish language input for each of the Mexican-American children. This omission could obscure any teacher relationship, especially when measured in terms of English reading ability. However, the insignificant effects of schools on these children make it difficult to argue against community control plans for this community.

A large caveat is needed at this point. The only measure of output used in this paper has been achievement test scores. This seems to be very important in terms of further education as that builds upon this foundation. However, this is probably not the only output in schools. In particular, teachers of Mexican-American children may spend a large proportion of their time on socialization aspects of education, e.g., discussing the American heritage or accepted behavioral patterns. This type of instruction by teachers, although somewhat improbable, could lead to the results of Equation 4.

There seem to be a number of directions in which one could proceed at this point. It is obvious that more information about the different dimensions of teacher quality is needed. One must be able to break down the verbal facility measure used in this paper. At the same time it is necessary to develop a model in terms of attributes which the administrator can purchase. While

some analysis, particularly that of Levin, suggests that schools implicitly buy attributes such as teacher verbal facility, buying these through a scale in terms of experience and education cannot help but be inefficient.* Further, it is evident through comparing verbal scores for teachers with national norms that present salary schedules do not attract the best college graduates into teaching. However, more information is needed about the supply schedules for specific teacher attributes.

At the same time it appears to be very important to expand the measures of output. Achievement test scores certainly do not reflect all dimensions of educational output. The relationship among different outputs of education is very imperfectly understood at this point.

Finally, it is important to broaden the California type sample. It is necessary to develop refined samples over a wide range of experiences. This includes matching students with specific inputs. It is necessary to look at different grades and different school systems. Further, the necessity of refining our measures of teachers is obvious.

---

*See Recruiting Teachers.

APPENDIX A

COMPLETE MULTISYSTEM SCHOOL MODELS (Verbal ability)
(log-log models)

| Variable | WHITE Coefficient (t statistic) | BLACK Coefficient (t statistic) |
|---|---|---|
| Central City: = 1 if cc<br>= 0 otherwise | -.025<br>(-4.1) | -.042<br>(-2.5) |
| Goods in home (average number with auto, TV, refrigerator, record player and phone) | .599<br>(10.4) | .662<br>(7.9) |
| Father's education (years) | .133<br>(4.4) | .022<br>(.4) |
| People in home | -.049<br>(1.8) | -.177<br>(-3.0) |
| % who attended nursery school | .015<br>(4.0) | |
| % student out migration during past year | -.005<br>(-1.8) | |
| % who wish to finish high school or more | .319<br>(4.8) | .590<br>(5.5) |
| % who feel they don't have much chance for success | -.027<br>(5.9) | -.028<br>(-2.3) |
| Racial concentration: = % Negro if between 45 and 75 percent<br>= 0 otherwise | | -.011<br>(-2.5) |
| Racial concentration: = % Negro if greater than 75 percent<br>= 0 otherwise | -.036<br>(-3.3) | -.006<br>(-1.3) |
| % with nonwhite teacher during the past year | -.024<br>(-7.1) | -.026<br>(-1.7) |
| Average score on teacher verbal test | .117<br>(2.2) | .178<br>(2.0) |
| Average years of teaching experience | .020<br>(3.2) | .045<br>(2.6) |

GSA DC 70-8812